

MODIS Level 2 IMAPP Atmospheric Science Processing Algorithm

IMAPP_SPA

General

The NASA Goddard Space Flight Center's (GSFC) Direct Readout Laboratory (DRL), Code 606.3 developed this wrapper software for the National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP) In-Situ Ground System (NISGS) and the International Polar Orbiter Processing Package (IPOPP).

Users must agree to all terms and conditions in the Software Usage Agreement on the DRL Web Portal before downloading this software.

Software and documentation published on the DRL Web Portal may occasionally be updated or modified. The most current versions of DRL software are available at the DRL Web Portal:

<http://www.directreadout.sci.gsfc.nasa.gov>

Questions relating to the contents or status of this software and its documentation should be addressed to the DRL via the Contact Us mechanism at the DRL Web Portal:

<http://directreadout.sci.gsfc.nasa.gov/index.cfm?section=contact%20usAlgorithm>

Algorithm Wrapper Concept

The DRL has developed an algorithm wrapper to provide a common command and execution interface to encapsulate multi-discipline, multi-mission science processing algorithms. The wrapper also provides a structured, standardized technique for packaging new or updated algorithms with minimal effort.

A Science Processing Algorithm (SPA) is defined as a wrapper and its contained algorithm. SPAs will function in a standalone, cross-platform environment to serve the needs of the broad Direct Readout community. Detailed information about SPAs and other DRL technologies is available at:

<http://directreadout.sci.gsfc.nasa.gov/index.cfm?section=technology>

Software Description

This DRL software package contains the MODIS IMAPP_SPA (International MODIS/AIRS Processing Package SPA). The IMAPP_SPA processes MODIS Aqua and Terra Level 1B Direct Broadcast (DB) data into four Level 2 MODIS atmospheric products: Cloudmask (MOD35); Cloudtop Properties and Cloud Phase (a portion of MOD06); Atmospheric Profiles (MOD07); and Aerosol (MOD04).

Software Version

Version 1 of the DRL algorithm wrapper was used to package the SPA described in this document. The SPA uses IMAPP Version 2.0 (August 28, 2007) processing code to generate MODIS Level 2 Atmospheric products, and incorporates MODIS Destripe Direct Broadcast Software Version 1.0 (September 24, 2008) destriping code.

Enhancements to this SPA include:

- a) a destriping capability to remove artificial stripes that appear in the MODIS L1B 1km infrared bands due to detector-to-detector variations and mirror side effects;
- b) grib2 ancillary data capability;
- c) improved ancillary file logic in IPOPP mode to provide the highest-quality science data for real-time applications;
- d) installable as an International Polar Orbiter Processing Package (IPOPP) plug-in (IPOPP Version 1.6a or later).

Credits

The IMAPP and MODIS Destripe Direct Broadcast software packages were provided to the DRL by the Space Science and Engineering Center (SSEC), University of Wisconsin-Madison.

Prerequisites

To run this package, you must have the Java Development Kit (JDK) or Java Runtime Engine (JRE) (Java 1.5 or higher) installed on your computer, and the bin directory of your Java installation in your PATH environment variable.

Program Inputs and Outputs

Inputs to the IMAPP_SPA are as follows:

- a) the MODIS 1km, half km, and quarter km L1B Calibrated Geolocated Radiances HDF products;
- b) the MODIS Geolocation HDF product;
- c) ancillary files for snow/ice extent, sea-ice concentration, sea-surface temperature, ozone, and weather.

The IMAPP_SPA outputs the following MODIS Level 2 atmospheric products:

- a) Cloudmask (MOD35);

- b) Cloudtop Properties and Cloud Phase (a portion of MOD06);
- c) Atmospheric Profiles (MOD07);
- d) Aerosol (MOD04).

Installation and Configuration

This section contains instructions for installing an SPA in a standalone configuration. SPAs may also be installed dynamically into an IPOPP framework; instructions for this type of installation are contained in the IPOPP User's Guide.

Download the IMAPP2.0_SPA.tar.gz and IMAPP2.0_SPA_testdata.tar.gz (optional) files into the same directory.

Decompress and un-archive the IMAPP2.0_SPA.tar.gz and IMAPP2.0_SPA_testdata.tar.gz (optional) files:

```
$ tar -xzf IMAPP2.0_SPA.tar.gz
$ tar -xzf IMAPP2.0_SPA_testdata.tar.gz
```

This will create the following subdirectories:

```
SPA
  IMAPP
    algorithm
    ancillary
    wrapper
    stations
    testscripts
    testdata
```

For convenience, statically pre-compiled versions of the required executables are included within the SPA (within SPA/IMAPP/algorithm/bin). The binaries should work on most Linux OS/platforms. The software is supported on the following Linux operating systems using the g77 gnu (gcc 3.2.1) compiler:

- a) Red Hat Linux x86 7.2;
- b) Red Hat Enterprise Linux 3 x86_64;
- c) Linux CentOS 4.2 x86_64 on Sun Opteron platforms.

The binaries have also been tested successfully on Fedora Core 4 and 6 platforms. If you receive an error message while running the testscripts (refer to the next

section, "Software Package Testing and Validation"), you may need to compile the software for your platform/OS combination. Refer to the Appendix for instructions on recompiling.

Software Package Testing and Validation

The testscripts subdirectory contains test scripts that can be used to check that your current installation of the SPA is working properly, as described below. Note that the optional IMAPP2.0_SPA_testdata.tar.gz file is required to execute these testing procedures.

Step 1: cd into the testscripts directory

Step 2: There are two scripts inside the testscripts directory: 'run-imapp' and 'run-imapp-autodownload'. The 'run-imapp' script uses input L1B files and ancillary files in the testdata/input directory to create the atmospheric products. The 'run-imapp-autodownload' uses the same L1B files; however it automatically downloads the appropriate ancillary files needed to process the L1B files. Run each one of them as follows:

```
$/run-imapp
```

```
$/run-imapp-autodownload
```

A successful execution usually takes some time (between 10 and 20 minutes, depending on the speed of your computer), so if the execution seems to get stuck with an "Awaiting Completion" message, do not become impatient. If everything is working properly, the scripts will terminate with a message such as:

Output modis.cloudmask is

/home/IPOPP/SPA/IMAPP/testdata/output/mod35.07041190603.hdf

Output modis.cloudtop is

/home/IPOPP/SPA/IMAPP/testdata/output/mod06.07041190603.hdf

Output modis.aerosols is

/home/IPOPP/SPA/IMAPP/testdata/output/mod04.07041190603.hdf

Output modis.atmprofile is

/home/IPOPP/SPA/IMAPP/testdata/output/mod07.07041190603.hdf

You can cd to the output directory (SPA/IMAPP/testdata/output) to verify that the science products exist. If the products exist, then the SPA works perfectly. Test output product(s) are available for comparison in the testdata/output directory. These test output product(s) were generated on a 64-bit PC architecture computer running Linux Fedora Core 4. Use a comparison utility (such as diff, hdiff, etc.) to compare your output product(s) to those provided in the testdata/output directory. If there is a problem and the code terminates abnormally, the problem can be

identified using the log files. Log files are generated automatically within the directory used for execution. They start with stdfile* and errfile*. Please report any errors that cannot be fixed to the DRL. You can delete the log files after execution. Test output product(s) are available for comparison in the testdata/output directory.

Program Operation

In order to run the package using your own input data, you can either use the 'run' scripts within the wrapper subdirectories, or modify the test scripts within the testscripts subdirectory.

To Use the Run Scripts

Identify the 'run' scripts: The wrapper/IMAPP subdirectory contains an executable called 'run'. Execute 'run' to execute the IMAPP_SPA and create MODIS atmospheric products. Note that to execute 'run', you need to have java on your path.

Specify input parameters using <label value> pairs: To execute the 'run' scripts, you must supply the required input and output parameters. Input and output parameters are usually file paths or other values (e.g., platform name or scan time). Each parameter is specified on the command line by a <label value> pair. Labels are simply predefined names for parameters. Each label must be followed by its actual value. Each process has its own set of <label value> pairs that must be specified in order for it to execute. Some of these pairs are optional, meaning the process would still be able to execute even if that parameter is not supplied. There are three types of <label value> pairs that the MODIS IMAPP_SPA uses, as follows:

- a) Input file label/values. These are input file paths. Values are absolute or relative paths to the corresponding input file.
- b) Parameter label/values. These are parameters that need to be passed onto the SPA (e.g., platform name or scan time).
- c) Output file labels. These are output files that are produced by the SPA. Values are the relative/absolute paths of the files you want to generate.

The following tables contains labels, and their descriptions, required by the MODIS IMAPP_SPA.

Input File Labels	Description	Data Sources
modis.mxd021km	MODIS 1km L1B Calibrated Geolocated Radiances HDF file (MOD021KM, MYD021KM)	DRL ftp site for real-time datasets over the eastern US region: Terra ftp://is.sci.gsfc.nasa.gov/gsfcddata/terra/modis/level1/ Aqua ftp://is.sci.gsfc.nasa.gov/gsfcddata/aqua/modis/level1/ Datasets from your Direct Readout station
modis.mxd02hkm	MODIS 500m L1B Calibrated Geolocated Radiances HDF file (MOD02HKM, MYD02HKM)	
modis.mxd02qkm	MODIS 250m L1B Calibrated Geolocated Radiances HDF file (MOD02QKM, MYD02QKM)	
modis.mxd03	MODIS Geolocation hdf file (MOD03, MYD03)	
ssmi_nise (optional)	NSIDC NISE (Near-real time snow and ice extent) (1 degree, global, daily).	DRL website for data from the last few days: ftp://is.sci.gsfc.nasa.gov/ancillary/temporal/global/ Archived ancillary data can be found at: ftp://ftp.ssec.wisc.edu/pub/eosdb/ancillary/
ssmi_seaice (optional)	National Centers for Environmental Prediction (NCEP) sea ice concentration (1 degree, global, daily)	
ncep_met (optional)	NCEP Numerical Weather Prediction GRIdded Binary (GRIB) File (6 hourly, 1 degree, global), which can be either Global Data Assimilation System (GDAS1) analysis fields or Global Model Forecast Fields (GFS) .	
noaa_oisst (optional)	National Oceanic and Atmospheric Administration (NOAA) optimum Interpolation Sea Surface Temperature (SST) (oisst- 1 degree, global, weekly)	

Input File Labels	Description	Data Sources
noaa_toast (optional)	NCEP Total Ozone Analysis using SBUV/2 and TOV) TOAST (daily, global)	
Parameter Labels	Description	
platform	'aqua' or 'terra'	
scantime	The start time of the L1B swath in yyyydddhhmm format. Note that ddd refers to the day of the year. If your input L1B files follows the standard DAAC L1B file naming convention (e.g., MYD021KM.Ayyyydddhhmmss.xxxxxxx.hdf), this information can be found in the file name itself.	
destp (optional)	The destriping capability is enabled by default. Destriping removes artificial stripes that appear in the MODIS L1B 1km infrared bands due to detector-to-detector variations and mirror side effects. To disable destriping, add the 'destp' parameter with the value "DestripeOff".	
gribtype (optional)	The grib1 ancillary data are used by default. To use grib2 ancillary data, add the 'gribtype' parameter with the value "2".	
Output File Labels	Description	
modis.cloudmask (optional)	Cloudmask Product output HDF file	
modis.aerosols (optional)	Aerosol Product output HDF file	
modis.cloudtop (optional)	Cloudtop Product output HDF file	
modis.atmprofile (optional)	Atmospheric Profiles output HDF file	

NOTES:

- Selective product processing:** The IMAPP_SPA will produce only those atmospheric products whose output labels were specified on the command line. For example, specifying only modis.cloudmask and modis.aerosol labels on the command line will result in generation of only the cloudmask and aerosol products. Output file format information for each product can be found in /SPA/IMAPP/algorithm/doc.
- Ancillary input files:** Specification of any ancillary input file is optional. If you do not specify a particular ancillary input label, an appropriate ancillary file will be downloaded automatically for processing. Make sure that you have an Internet connection if you do not specify the ancillary files on the command line. We recommend that you use the automatic download feature by not specifying ancillaries on the command line. You may also download the appropriate ancillary files from the data sources provided in the table before

execution. Recommendations on choosing ancillary input files for a particular L1B granule are as follows:

- The dates for the National Snow and Ice Data Center (NSIDC) near-real time snow and ice and NCEP sea ice datasets should be as close as possible to the dates of the L1B granules. The date of the OISST weekly ancillary file should also correspond as closely as possible to the L1B scan time. The dates can be identified from the filenames.
- The dates of the TOAST ozone data should be as close as possible to the dates of the L1B granules. The dates can be identified from the filenames. For processing L1B granules created prior to March 31, 2005, TIROS Operational Vertical Sounder (TOVS) Ozone ancillary data should be used. They are available in the ftp archive site provided in the table above.
- GDAS files are produced every 6 hours at 00, 06, 12, and 18 UTC daily. The time, date and hour of the GDAS files in grib 1 format can be found in the filename. For example, gdas1.PGrbF00.070210.18z corresponds to February 10, 2007, 1800 UTC. When choosing the GDAS ancillary file, choose one which is closer in time rather than the date. For example if you have an L1B granule at 1700 UTC, the GDAS file for 1800 hours on the same day would be the best match. However, if that is not available, it would be preferable to use the GDAS file corresponding to 1800 UTC for the day before rather than the GDAS file at 600 UTC for the same day. Real time data sets may also use the Global Model Forecast Fields (GFS) instead of GDAS files, since GDAS files may not be available. Recent GFS data in grib1 format from the last few days can be acquired from the DRL ftp site: <ftp://is.sci.gsfc.nasa.gov/ancillary/temporal/global/gfs/> The naming conventions for grib1 gfs files are gfs.thh.yymmdd.pgrbfxx (Here yymmddd and hh represent analysis time, and xx represents forecast time step). If you choose GFS data as input, you should attempt to use a forecast time step that is closest to the analysis time. For example, if your data time is 15 UTC, you should try to use the 3 hour forecast field from the 12 UTC model run, instead of the 9 hour forecast field from the 06 UTC run.

Execute the 'runs': The following is an example of a command line to run the IMAPP_SPA from the testscripts subdirectory:

```
$ ../wrapper/IMAPP/run \
  modis.mxd021km ../testdata/input/MYD021KM.07041190603.hdf \
  modis.mxd02hkm ../testdata/input/MYD02HKM.07041190603.hdf \
  modis.mxd02qkm ../testdata/input/MYD02QKM.07041190603.hdf \
  modis.mxd03 ../testdata/input/MYD03.07041190603.hdf \
  ssmi_nise ../testdata/input/NISE_SSMIF13_20070210.HDFEOS \
  ssmi_seaice ../testdata/input/eng.070210 \
  ncep_met ../testdata/input/gdas1.PGrbF00.070210.18z \
  noaa_oisst ../testdata/input/oisst.20070207 \
  noaa_toast ../testdata/input/TOAST16_070210.GRB \
  platform aqua \
  scantime 20070411906 \
  modis.cloudmask ../testdata/output/mod35.07041190603.hdf
  modis.aerosols ../testdata/output/mod04.07041190603.hdf
```

A successful execution of 'run' usually takes some time (around 10-20 minutes, depending on the speed of your computer), so if the execution seems to get stuck with an "Awaiting Completion" message, do not become impatient. If execution fails, you will see an error message indicating the cause of failure (e.g., a file cannot be found, or a label cannot be recognized). Correct it and run again. If the problem has some other cause, it can be identified using the log files. Log files are automatically generated within the directory used for execution. They start with stdfile* and errfile* and can be deleted after execution. Please report any errors that cannot be fixed to the DRL. The 'run' can be executed from any directory the user chooses. This can be done by prefixing it with the file path for the 'run' script.

To Use the Script in the Testscripts Directory

One simple way to run the algorithms from any directory of your choice using your own data is to copy the test scripts from the testscripts directory to the selected directory. Change the values of the variables WRAPPERHOME, L1HOME, ANCHOME and OUTPUTDIR to reflect the file paths of the wrapper directories and the input/output directory locations. Then modify the input/output file name variables. Run the scripts to process your data.

Appendix

Instructions for Compiling the IMAPP_SPA Software

If you cannot use the pre-compiled binaries located in the SPA/IMAPP/algorithm/bin directory, you will need to compile the code manually. A script to manually compile all of the executables is included in this software package (IMAPP/algorithm/DRLshellscripts/make_imapp_executables.sh). The script will attempt to compile all of the source code located within the SPA/IMAPP/algorithm/src directory and copy the executables into the bin directory. The script uses the g77 compiler and is set up to try and compile on IRIX64, SunOS, AIX, Linux and HP-UX operating systems. You will need to set the HDF_HOME and MODIS_L2_HOME environmental variables prior to executing the script. This can be done manually within the script file, or through sourcing one of the environmental variable scripts (environmental variable scripts for bash and csh are included in the SPA/IMAPP/algorithm/env directory) after editing the latter variables to the correct values. MODIS_L2_HOME must be set to the SPA/IMAPP/algorithm directory, while HDF_HOME must be set to your hdf installation directory. You will need to install the HDF4.2r1 libraries on your system. HDF4.2r1 can be obtained in pre-compiled binary form from: <http://www.hdfgroup.org/release4/obtain.html>. Make sure you get the SZIP, ZLIB and JPEG libraries also, and place all of the library and include files into the main hdf/lib and hdf/include directories.

If you wish to compile only a part of the source code, cd into the individual product src or cmd directory, type "make", and copy the executable into the SPA/IMAPP/algorithm/bin directory. You may need to set the \$HDF_HOME environmental variable first. This can be done manually, within the make file, or through sourcing one of the environmental variable scripts.

Other required binaries include hdp, ncftpget, ncdump, and wgrib. All of these are located within the SPA/IMAPP/algorithm/bin directory. All have been compiled statically and tested on the supported platforms. If you receive an error message when executing hdp or ncdump, you must obtain the correct binaries for your platform. These utilities are a part of the HDF distribution, and pre-compiled binaries can be acquired from: ftp://ftp.ncsa.uiuc.edu/HDF/HDF/HDF_Current/bin. The ncftpget utility is also available for different platforms via download from: <http://www.ncftp.com/ncftp/>. Finally, wgrib, the utility to unpack fields from the ancillary GRIB files can be acquired from: <http://www.cpc.ncep.noaa.gov/products/wesley/wgrib.html>